Position-Dependence of Stabilizing Polar Interactions of Asparagine in Transmembrane Helical Bundles†

James D. Lear,*,‡ Holly Gratkowski,‡ Larisa Adamian,§ Jie Liang,*,§ and William F. DeGrado*,‡

Department of Biochemistry and Biophysics, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6059, and Department of Bioengineering, University of Illinois at Chicago, 851 South Morgan Street, MC-063, Chicago, Illinois 60607

Received September 5, 2002; Revised Manuscript Received February 14, 2003

ABSTRACT: Recent studies with model peptides and statistical analyses of the crystal structures of membrane proteins have shown that buried polar interactions contribute significantly to the stabilization of the three-dimensional structures of membrane proteins. Here, we probe how the location of these polar groups along the transmembrane helices affect their free energies of interaction. Asn residues were placed singly and in pairs at three positions within a model transmembrane helix, which had previously been shown to support the formation of trimers in micelles. The model helix was designed to form a transmembrane coiled coil, with Val side chains at the "a" positions of the heptad repeat. Variants of this peptide were prepared in which an Asn residue was introduced at one or more of the "a" positions, and their free energies of association were determined by analytical ultracentrifugation. When placed near the middle of the transmembrane helix, the formation of trimers was stabilized by at least −2.0 kcal/mol per Asn side chain. When the Asn was placed at the interface between the hydrophobic and polar regions of the peptide, the substitution was neither stabilizing nor destabilizing (0.0 ± 0.5 kcal/mol of monomer). Finally, it has previously been shown that a Val-for-Asn mutation in a water-soluble coiled coil destabilizes the structure by approximately 1.5 kcal/mol of monomer [Acharya, A., et al. (2002) *Biochemistry 41*, 14122−14131]. Thus, the headgroup region of a micelle appears to have a conformational impact intermediate between that of bulk water and the apolar region of micelle. A similarly large dependence on the location of the polar residues was found in a statistical survey of helical transmembrane proteins. The tendency of different types of residues to be buried in the interiors versus being exposed to lipids was analyzed. Asn and Gln show a very strong tendency to be buried when they are located near the middle of a transmembrane helix. However, when placed near the ends of transmembrane helices, they show little preference for the surface versus the interior of the protein. These data show that Asn side chains within the apolar region of the transmembrane helix provide a significantly larger driving force for association than Asn residues near the apolar/polar interface. Thus, although polar interactions are able to strongly stabilize the folding of membrane proteins, the energetics of association depend on their location within the hydrophobic region of a transmembrane helix.

Our understanding of the features that stabilize the folds of membrane proteins is beginning to emerge through the study of natural proteins (*1*−*3*) and designed systems (*4*, *5*). Much information has been gleaned from the study of crystal structures of helical membrane proteins (*6*−*8*), although the size of the database is still small. For example, a recent statistical study (*7*) revealed that interhelical hydrogen bonds are particularly important for the stabilization of helix−helix interactions in membrane proteins; most transmembrane helices in polytopic membrane proteins have at least one interhelical hydrogen bond, and hydrogen-bonded transmembrane helices tend to interact more extensively than non-hydrogen-bonded helices. Similarly, experimental studies involving simple transmembrane helices have shown that a single polar residue, such as Asn, Asp, Glu, Gln, or His, is sufficient to mediate homooligomerization of the helix in biological membranes and micelles (*9*−*12*).

Previously, we developed a model peptide system for assessing the features influencing the association of transmembrane peptides. The peptide is based on GCN4-P1 (*10*), corresponding to a coiled coil in the water-soluble protein GCN4. The wild-type GCN4-P1 peptide forms dimers, while many variants of this peptide form trimeric structures (*13*, *14*). As in GCN4-P1, our membrane-solubilized peptide has an approximate seven-residue repeat, with Leu residues occupying the "d" positions, and Val residues at all but one "a" position. Randomly selected hydrophobic residues were substituted into the remaining positions over a 20-residue stretch, which was intended to form a transmembrane helix. The membrane-soluble version of GCN4-P1 associates in micelles and bilayers (*10*). In micelles, it exists in a monomer−dimer−trimer equilibrium, the cooperativity of

† Supported by NIH GM60610.

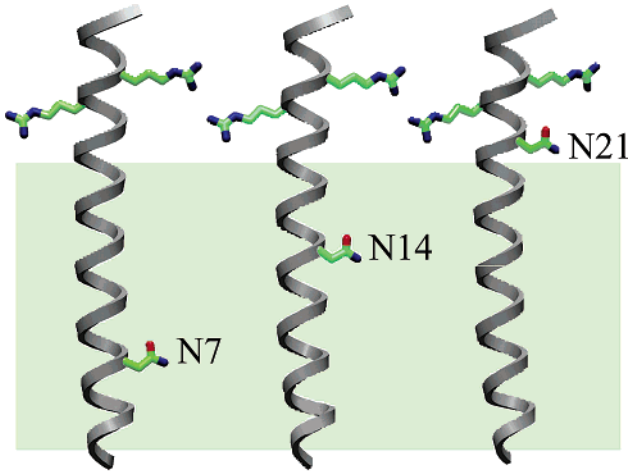
* To whom correspondences should be addressed. E-mail: wdegrado@mail.med.upenn.edu. Tel: (215) 898-4590; fax: (215) 573-7229.

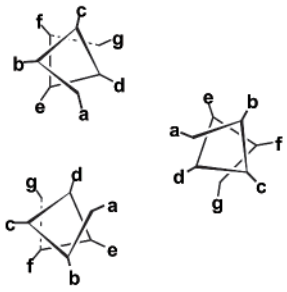
‡ University of Pennsylvania.

§ University of Illinois at Chicago.

[1] Abbreviations: SDS, sodium dodecyl sulfate; NBD, 7-nitrobenz-2-oxa-1,3-diazole; C14-betaine, *N*-tetradecyl-*N*,*N*-dimethyl-3-ammonio-1-propanesulfonate.

10.1021/bi020573j CCC: $25.00
Published on Web 05/08/2003






```
          d    a  d    a  d    a  d    a
N7      BQLIAA NLLLIAV VLILLIA VARLRYL VG
N14     BQLIAA VLLLIAV NLILLIA VARLRYL VG
N21     BQLIAA VLLLIAV VLILLIA NARLRYL VG
N7-N21  BQLIAA NLLLIAV VLILLIA NARLRYL VG
N14-N21 BQLIAA VLLLIAV NLILLIA NARLRYL VG
A7-N14  BQLIAA ALLLIAV NLILLIA VARLRYL VG
```

FIGURE 1: Sequence of peptides and the location of Asn side chains within the hydrophobic region. In the top panel, the transmembrane region of the peptides is indicated by the shaded box. The helical models show an Asn placed at N7, N14, and N21 from left to right. The C-terminal polar residues include two Arg side chains that are shown near the top of the figure. Atoms are colored blue for nitrogen, red for oxygen, and green for carbon. The middle panel illustrates a schematic diagram showing the heptad repeat of a coiled coil. The peptide sequences are depicted in the lower panel.

assembly depending on the type of micelle in which it is studied (*19*). Independently and simultaneously, Engelman and co-workers showed that a similar peptide associated in SDS micelles and biological membranes (*12*). In both studies, the single Asn side chain proved critical to the association; the replacement of this Asn with a Val side chain led to a peptide with a very weak tendency to form oligomers. This finding contrasts with results from the water-soluble GCN4-P1, in which an Asn-to-Val mutation stabilized the formation of dimers and trimers (*14*).

Here we quantitatively determine how the position of an Asn side chain in different regions of the transmembrane helix affects its free energy of association. An Asn side chain was individually placed at each of the "a" positions of the peptide, ensuring that the helix packing interactions would remain approximately the same in each case. We also examine the effects of simultaneously placing an Asn side chain at two "a" positions of the peptide. These studies show that when placed within the hydrophobic region of the transmembrane helix, an Asn side chain provides a large driving force for association of the helices, and that this effect is attenuated as the Asn is moved closer to the polar end.

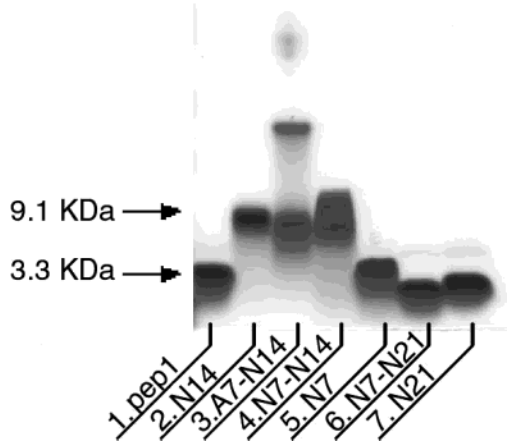


FIGURE 2: SDS−PAGE of peptides. Five microliters of a 1 mg/mL sample was loaded in each well as described in the methods.

RESULTS

Design of Peptides. We prepared a series of three peptides in which a single Asn residue was placed at each possible "a" position within the transmembrane sequence of our previously described membrane-soluble version of GCN4-P1 (MS1). Here, we adopt a new nomenclature to facilitate discussion of these and other variants (Figure 1). The parent peptide is designated pep1; it contains Val at each "a" position. N7, N14, and N21 have Asn side chains at a single "a" position within the transmembrane helix (Figure 1). Asn7 is located within the transmembrane region, five residues from the closest polar side chain, and Asn14 is located near the center of the transmembrane helix. By comparison, Asn21 is located at the C-terminal end of the transmembrane helix; there is only one intervening Ala between N21 and the first charged residue (Figure 1). Further, to determine the effects of introducing two Asn side chains, we prepared N7-N14, in which both Asn residues lie within the hydrophobic region of the transmembrane. We also prepared N7-N21, in which one side chain lies within the hydrophobic region and the other lies near the interfacial region.

To determine the role of a buried Val side chain at an "a" position, A7-N14 was also synthesized, in which Val7 is changed to Ala and Val14 is changed to Asn. In a water-soluble coiled coil, the replacement of Val7 with a smaller Ala side chain would be expected to destabilize the association, due to the decreased hydrophobic driving force. It is more difficult to anticipate the effect of such a mutation in a membrane-soluble protein. On one hand, small residues are frequently observed at helix−helix interfaces of membrane proteins (*6*−*8*, *15*, *16*) where they presumably stabilize folding, while on the other hand one might expect the smaller Ala side chain would lead to cavity formation, which would destabilize the protein.

SDS Polyacrylamide Gel Electrophoresis (SDS−PAGE). SDS−PAGE provides a very rapid method to assess the association of the peptides in micelles (Figure 2). Previously, we demonstrated that N14 migrated at the position expected for a trimer (*10*), while the parent peptide with Val at each

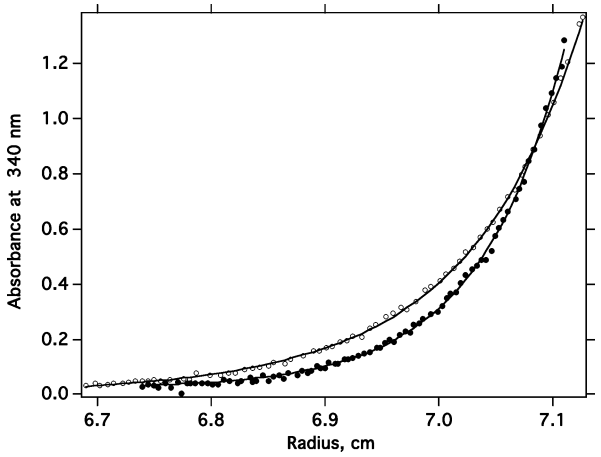


FIGURE 3: Comparison of analytical ultracentrifugation data for two peptides with different free energies of association. The data were collected at 50 000 rpm. for A7-N14 (open circles) and N7-N14 (solid circles). A steeper absorbance gradient is observed for N7-N14, which has a more favorable free energy of association (Table 1). Lines are generated from single fits to the data for a monomer−dimer trimer model for A7N14, and to a single species trimer for N7N14.

"a" position migrated as a monomer. Here, we find that the two double mutants with Asn at position14, N7-N14 and A7-N14 also migrated at approximately the position expected for a trimer. Significantly, A7-N14 also showed bands at higher molecular weight, suggesting that the introduction of Ala has resulted in a decrease in the conformational specificity of the peptide. All of the peptides that lacked Asn at the most central region of the transmembrane helix showed a single major band at a position near that expected for the monomer.

Analytical Ultracentrifugation. While SDS gel electrophoresis provides a rapid test of the association of transmembrane helices, it is not performed under equilibrium conditions and hence does not provide quantitative information concerning the thermodynamics of association. We therefore performed analytical ultracentrifugation (Figure 3) of the peptides dissolved in micelles formed from the zwitterionic detergent, C14-betaine. This detergent was chosen because it has a density which can be matched by using H_2O/D_2O buffer mixtures, and it has been shown to increase the cooperativity of association of N14 (*19*). Figure 3 illustrates analytical ultracentrifugation data for A7-N14 versus N7-N14 as representative examples of peptides with a single versus two Asn side chains. The addition of a second Asn in N7-N14 leads to a steeper concentration gradient, which is indicative of a greater degree of association. Complete data for all of the variants are provided in Supporting Information.

While a fully cooperative monomer−trimer equilibrium adequately fit the sedimentation curves of several of the peptides, others required a noncooperative monomer−dimer−trimer scheme to provide an adequate fit. Thus, the sedimentation curves for all the peptides were fit to a monomer−dimer−trimer equilibrium, which provided a measure of the cooperativity of assembly, as well as a realistic estimate of the uncertainties in the trimerization equilibrium constants. To obtain an accurate determination of the thermodynamics of association, all peptides were centrifuged at two or more concentrations and rotor speeds,

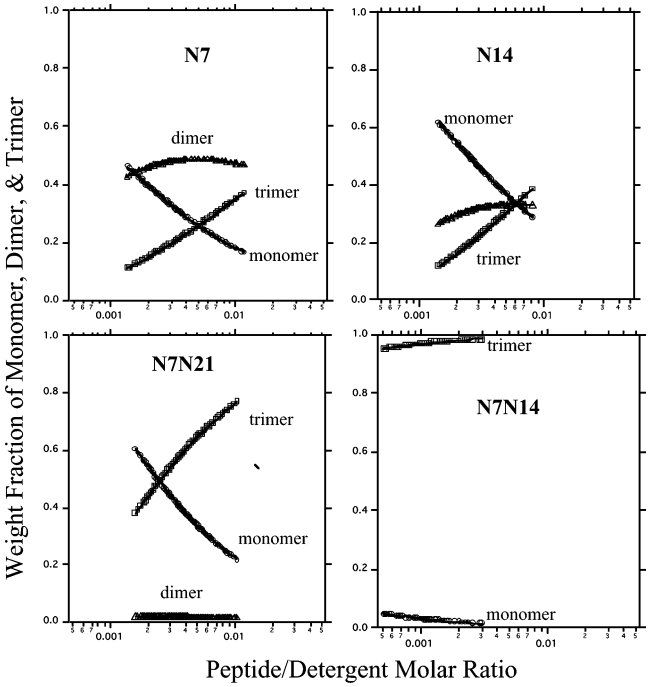


FIGURE 4: Weight fraction of various molecular species as a function of peptide/detergent ratio for N7, N14, N7-N14, and N7-N21. The curves were generated from the best-fit dissociation constants for the dimerization and trimerization interactions, and are shown over the experimentally observed range of peptide/detergent ratios. Analysis of N7-N14 indicated that the concentration of the dimeric intermediate for the data was undetectably low and hence is not drawn in the figure.

Table 1: Free Energies of Trimerization[a]

peptide	ΔG_{trimer} (kcal/mol)
pep1	> -5.0
N7	-7.3 ± 1
N14	-6.7 ± 0.6
N21	>-5.0
N7-N14	-13.6 (-1.9 + n.d.)[b]
N7-N21	-7.3 ± 1.4
A7-N14	-7.5 ± 0.2

[a] Energies are per mole of trimer and the errors represent 95% confidence intervals. [b] The confidence interval is asymmetric for this peptide; the lower limit (95%) is 1.9 kcal/mol, whereas the upper limit is not well defined.

and the dissociation constants for dimerization and trimerization were globally fit using all of the data for a given peptide (*9*). Full details of the method of error analysis and fits to the data are described in Supporting Information.

Figure 4 illustrates the deduced association curves obtained at the experimentally determined peptide/detergent ratios for N7, N14, N7-N14, and N21. It is interesting to note that the peptides with a single Asn side chain show a much lower cooperativity of assembly (i.e., they have higher concentrations of dimeric intermediates) than the corresponding peptides with two Asn side chains. Relatively high concentrations of a dimeric intermediate were also found for the single-Asn peptides A7-N14 and N21.

Table 1 provides the overall free energies of trimerization for each peptide, expressed as the free energy per mole of trimer. A unitary mole fraction (peptide/detergent ratio) was adopted as the standard state. It is useful to consider the source of the errors associated with these values of ΔG_{trim}.

First, for peptides that do not show a very cooperative assembly to form trimers (most notably N7, Figure 4), there is a large covariation in the monomer−dimer and the monomer−trimer equilibrium constants, giving rise to relatively large errors in the monomer−trimer equilibrium constants listed in Table 1. A second source of error arises for N21 and pep1, in which there is too little association of the monomers (see Supporting Information) to adequately define the final association state and the equilibrium constant. For this reason, free energies of association are not quoted for these two peptides. Finally N7-N14 is so stable that very little if any dissociation of the trimers is observed at the lowest experimentally accessible peptide/detergent ratio (Figure 4). Thus, we can only place a limit on the favorable free energy of association for this peptide as ≤ -13.6 kcal/mol. A sensitivity analysis indicated that other sources of error, such as the uncertainty in partial specific volume of the peptides, are relatively small compared to these effects.

The dependence of the ΔG_{trim} on the location of the Asn residue can be observed by comparing the free energy of trimerization for N7, N14, and N21 (Table 1). The Asn residue is within the apolar region of the transmembrane helix in N14 and N7, and these peptides show a quite favorable free energy of association. In N21, the Asn is closer to the interface between the transmembrane and charged region of the peptide; $\Delta\Delta G$ for trimerization of this peptide is at least 2 kcal/mol less favorable than for N7 and N14.

The double-Asn variants further support the conclusions from the single-Asn peptides and provide the first good estimates of how an Asn side chain affects the $\Delta\Delta G$ of association. In previous work, we used the parent, pep1 (with Val residues in each "a" position), as a standard to compare with N14 (*9*, *10*). However, the very low degree of association of this peptide led to a large uncertainty in its free energy of oligomerization. Thus, the values of $\Delta\Delta G$ reported in our earlier study should be considered as lower limits. Here, the availability of single and double mutants, both of which have better determined association curves, allows the determination of improved estimates of $\Delta\Delta G$. The values of $\Delta\Delta G$ for the various mutants are shown in Scheme 1.

Scheme 1

$$\Delta\Delta G(\text{V7} \rightarrow \text{N7}) = [\Delta G(\text{N7-N14}) - \Delta G(\text{N14})]/3 \leq -2.3 \text{ kcal/mol}$$

$$\Delta\Delta G(\text{V14} \rightarrow \text{N14}) = [\Delta G(\text{N7-N14}) - \Delta G(\text{N7})]/3 \leq -2.1 \text{ kcal/mol}$$

$$\Delta\Delta G(\text{V21} \rightarrow \text{N21}) = [\Delta G(\text{N7-N21}) - \Delta G(\text{N7})]/3 = 0 \text{ kcal/mol}$$

The values of $\Delta\Delta G$ reflect the energetic differences on a per-residue basis and have been divided by a factor of 3 to correct for the fact that there are three monomers per trimer. The free energy change associated with the mutation of Val7 and Val14 to Asn are the same within experimental error, and both values are significantly more favorable than -2.0 kcal/mol of monomer. In contrast, mutation of Val21 to N21 results in a value of $\Delta\Delta G$ of 0.0 ± 0.5 kcal/mol.

Table 1 also provides the free energy of association of A7-N14, in which Val7 has been converted to Ala. Although this peptide was found to form a minor population of higher order oligomers in SDS micelles by PAGE, its sedimentation curves in C14-betaine were very well described by a monomer−dimer−trimer equilibrium. Comparing the free energies of association of A7-N14 with N14 indicates that the replacement of Val with Ala results in no significant change in the free energy of association ($\Delta\Delta G = -0.3 \pm 0.3$ kcal/mol on a per-residue basis).

Distribution of Amino Acids within the Transmembrane Regions of Polytopic Membrane Proteins The above observations indicate that the energetics of polar interactions depend markedly on their location within the transmembrane helix. However, the experiments were conducted with model peptides in micelles, which may not be fully representative of the situation in phospholipid bilayers, or more importantly, cellular membranes. Thus, to determine whether these results are representative of the fundamental forces responsible for the structure and evolution of the sequences of membrane proteins, we determined the frequencies of occurrence of various amino acid types on the lipid-accessible surface versus the interior of the protein.

A spherical probe was used to determine the accessibility of transmembrane helices to lipid molecules. It is unclear which probe radius should be used for the measurement of lipid-accessibility; we therefore performed calculations with a probe radius of 2.5 and 5.0 Å. The calculations were conducted on a database of transmembrane helices, which included the partially polar interfacial helical residues as well as the central predominantly nonpolar transmembrane region. We compared the probe-accessibility for residues within five residues of either end of a transmembrane helix versus the remaining positions near the middle of the helices (Table 2). The fractional surface-accessibility for a specific amino acid type was then compared to the overall population of all the amino acids, leading to lipid surface propensities, P_{surf}, for residues in either the end or the middle of the transmembrane segment. Qualitatively similar behavior was observed with a 2.5 and 5.0 Å radius, although the larger probe size tended to better discriminate between the preferences for buried versus exposed side chains.

Figure 5 illustrates the surface propensities for the residues that have been placed at the helix−helix packing interface in the present study. In addition, we have included Ser and Thr, as examples of small polar amino acids with a significantly lower potential to mediate association of transmembrane helices than Asn. In this figure, the logarithm of P_{surf} is plotted versus the amino acid type, such that residues with a significant tendency to lie on the surface of the protein will have values larger than 0, and residues that tend to be buried will have values less than 0.

The hydrophobic Val and Leu side chains showed a modest deviation from the population average, and tend to occupy surface accessible positions, particularly near the center of the helices. Ala has a modest but significant tendency to be buried in the protein; this observation is consistent with the previously observed tendencies of small side chains to lie at the helix−helix interfaces in membrane proteins (*8*, *15*, *16*). The value of P_{surf} for Ala depends little on its position within the helix. Ser and Thr show a similar position-independent tendency to be buried within transmembrane proteins. Serine has a slightly higher tendency to be buried than Thr, which is consistent with Ser being the

Table 2: Preferences for Amino Acids in the Middle and at the Ends of Transmembrane Helices

	end region					middle region				
		2.5 Å probe		5.0 Å probe			2.5 Å probe		5.0 Å probe	
amino acid	n_{total}[a]	n_{surf}[b]	P_{surf}[c]	n_{surf}	P_{surf}	n_{total}	n_{surf}	P_{surf}	n_{surf}	P_{surf}
Ala	106	46	**0.77**[d]	31	**0.72**	111	54	0.90	32	**0.74**
Val	84	48	1.0	39	1.14	128	87	**1.26**	68	**1.37**
Leu	151	100	**1.17**	78	**1.27**	174	126	**1.34**	95	**1.41**
Asn	16	6	0.67	6	0.92	14	4	0.53	1	**0.18**
Gln	17	10	1.0	8	1.16	6	0	**0**	0	**0**
Gln + Asn	33	16	0.86	14	1.04	20	4	**0.37**	1	**0.13**
Ser	37	12	**0.58**	8	**0.53**	55	19	**0.64**	9	**0.42**
Thr	47	22	0.83	13	**0.68**	69	28	**0.75**	20	**0.75**
Gln + Asn	33	16	0.86	14	1.04	20	4	**0.37**	1	**0.13**

[a] n_{total} is the total number of amino acids of a given type in the designated region (end or middle). [b] n_{surf} is the total number of amino acids of a given type that are accessible to a probe of a given radius (2.5 or 5.0 Å) within the designated region (end or middle). [c] P_{surf} is the ratio of the fraction of the surface accessible residues of a given type within a given region, relative to the overall fractional accessibility of all residues within the same region: $P_{surf} = [n_{surrf}(i)/n_{total}(i)]/[\Sigma\, n_{surrf}(i)/\Sigma\, n_{total}(i)]$ in which i represents a given amino acid type and the sum is taken over all 20 amino acids. [d] Values in bold deviate from the null hypothesis ($P_{surf} = 1.0$) at the 95% confidence level as assessed by the binomial distribution.

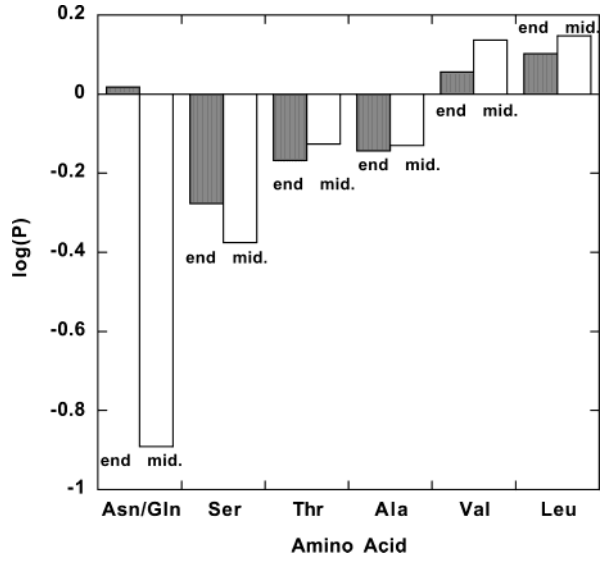


FIGURE 5: The logarithm of P_{surf} plotted versus amino acid. Values greater than zero indicate a propensity for residues to be exposed to lipid, and values less than zero indicate residues preferentially engage in interhelical interactions. The values of P_{surf} were calculated using a 5 Å probe to compute the probe-accessible surface.

more polar residue. Thr might be considered to act as an apolar amino acid in membrane proteins, because it is able to form an intrahelical hydrogen bond to a mainchain carbonyl preceding it by three or four residues in sequence. In this respect, it is therefore interesting to note that the isosteric amino acid, Val, behaves differently from Thr. As expected from the physicochemical properties of the side chains, Thr tends to be buried in membrane proteins while Val tends to be exposed to the phospholipids.

As compared to Thr and Ser, Asn and Gln show a much stronger tendency to be buried in the interior of membrane proteins when these side chains are placed within the apolar region of the bilayer. Table 2 provides data for these amino acids; because Gln and Asn are relatively rare in membrane proteins, the number of observations is not large. Thus, to improve the statistical significance of the data, we determined the propensity for the combined class of these two carboxamide-containing side chains. Unlike the other amino acids, P_{surf} for Asn/Gln depends quite markedly on the location of the side chains in the transmembrane helices. Near the ends of the helices, they partition equally between the interior and exterior of the protein. In contrast, near the center, these side chains are found almost exclusively in the interior of the protein, where they engage in multiple hydrogen-bonded and polar interactions. This behavior is quite similar to that observed experimentally, in which an Asn side chain near the end of the helix has a much lower propensity to engage in interhelical interactions than one near the center of the transmembrane segment. Also, previous experimental studies showed that Asn and Gln tend to mediate helix association more effectively than Ser or Thr, when these residues are placed near the center of the helix (*9*−*12*).

DISCUSSION

Previously, it was shown that the inclusion of polar amino acids into a transmembrane helix can provide a strong driving force for association. Here we determine how the placement of an Asn side chain at different locations within the hydrophobic region of the transmembrane helix affects the free energies of association. The heptad repeat of the parent peptide is an important aspect of the experimental design; in each variant, the Asn has a hydrophobic Leu side chain at the "d" positions immediately above and below it, ensuring that the packing interactions should be similar from peptide to peptide. Although it is possible that there are multiple effects associated with the location of the Asn within the peptide chain, it seems reasonable that the most important effect is its location within the hydrophobic domain.

The effects of the position of the Asn side chain depend somewhat on the nature of the micelle. In SDS micelles, the Asn side chain must be at position 14, near the center of the apolar stretch to support the formation of trimers. However, in zwitterionic betaine micelles, N7 and N14 showed similar free energies of trimerization, indicating that it was only necessary for the Asn to occur within the apolar region of the transmembrane helix.

These studies provide the first quantitative evaluation of the effects of location of polar groups on $\Delta\Delta G$ of their interaction. When placed near the center of a bilayer $\Delta\Delta G$ was favorable, and on the order of 2.0 kcal/mol. When placed at the interface between the polar and apolar regions, the value of $\Delta\Delta G$ was 0.0 ± 0.5 kcal/mol. By comparison, similar Ile-to-Asn mutations in water-soluble coiled coils are

generally destabilizing by approximately 1.5 kcal/mol of monomer (*20*, *21*). These differences can be rationalized in terms of the differences in positions of the Asn side chains in the monomeric forms of the peptides. In N7 and N14, the carboxamide-containing side chain would be expected to locate within the hydrophobic core of the betaine micelle, in N21 the Asn would locate to the less polar headgroup region, and in a water-soluble coiled coil the side chain would be fully exposed to water. Thus, as the polarity of the environment increases, the free energy associated with exposing the Asn side chain in the monomeric state becomes more favorable.

The results from the double mutations are also consistent with this expectation and indirectly support the hypothesis that the peptides adopt a coiled-coil like structure in the membrane. The more favorable free energy of trimerization associated with the addition of a second Asn side chain strongly suggests that the side chains can interact as in our guiding model. Engelman and co-workers have also demonstrated that the addition of a second Asn side chain can increase the degree of association of a model transmembrane helix in biological membranes as determined by the TOXCAT system (*12*). In this case, the transmembrane helix was formed from a poly-Leu structure, and it was fused to a globular protein that appeared to favor a dimeric rather than a trimeric conformation. Nevertheless, the net effect is similar to that observed in the current study.

The results of our statistical survey strongly support the conclusions from our experimental studies. The statistical study indicated that the hydrophobic side chains showed relatively weak tendencies to deviate from the random distributions expected from the population mean. Thus, Ala has a relatively weak preference to occupy the interior vs the exterior of a membrane protein, whereas Val and Leu have a weak tendency to be exposed to the lipid environment. This finding is consistent with our experimental finding that mutating Val7 to Ala had little effect on the formation of trimers. However, it should be remembered that the propensity to bury a given side chain clearly depends on its structural context (*15*, *16*).

The statistical survey also revealed that Ser and Thr behaved as polar amino acids in that they tend to be buried in helical proteins. However, when compared to the carboxamide-containing residues, Asn and Gln, the tendency of Ser and Thr to be buried was rather weak. Furthermore, the potential of Asn and Gln to form interactions within the protein, relative to being exposed on the surface, depended strongly on their position in the transmembrane segment. In both the experimental and statistical studies, a carboxamide-containing side chain shows much stronger interhelical interactions when it is located near the middle of the hydrophobic helix. Thus, when an Asn or Gln side chain occurs near the edge of the hydrophobic domain, it can be reasonably well solvated in the headgroup region of the bilayer or micelle. By contrast, when placed near the middle of a monomeric transmembrane helix, the carboxamide would be exposed to the apolar lipid hydrocarbon chains of the phospholipid bilayer or detergent micelle. This situation provides a strong driving force for folding into a well-packed structure that allows the polar side chains to engage in more stabilizing interhelical interactions.

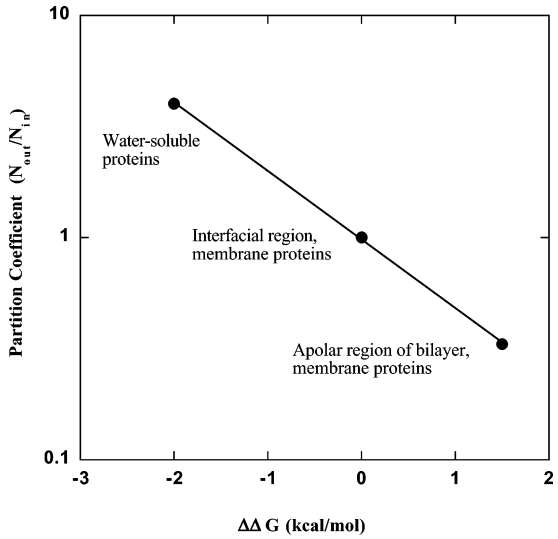


FIGURE 6: The partition coefficient for transfer of Asn+Gln from the inside to the outside of proteins as a function of the $\Delta\Delta G$ for an Asn-to-Val mutation. The partition coefficients are estimated from the number of buried versus nonburied Asn+Gln residues in water-soluble proteins,[22] from the interfacial region of membrane proteins (this work, 2.5 Å probe), and the central region of transmembrane helices in membrane proteins (this work, 2.5 Å probe). The value of $\Delta\Delta G$ for Asn-to-Val mutations of water-soluble proteins is taken as the average value obtained from from thermal, guanidine, and urea-unfolding studies of water-soluble coiled coils in which the active site a Val at an "a" position is converted to Asn (as described in ref 20 and 21). The value of $\Delta\Delta G$ for Asn-to-Val mutation in membrane proteins are from Scheme 1.

The excellent agreement between these experimental studies of model peptides dissolved in detergent micelles, and statistical studies of membrane proteins highlights the synergy between these two lines of investigation. It has long been recognized that the distribution of side chains in water-soluble proteins reflects evolutionary pressures relating to protein stability (*21*). It was early shown that the distributions of side chains in the interior versus the exterior of proteins are related to the free energies of transfer of the corresponding amino acids from the aqueous phase to solvents that mimic the interior of water-soluble proteins.

Our results extend this analysis to membrane proteins; in Figure 6 we examine the correlation between $\Delta\Delta G$ of mutation from this work—which provides information concerning stability in a micellar environment—with the partition coefficients obtained from analysis of membrane protein structures—which provides information concerning the forces stabilizing the folding of a protein in a membrane environment. The *y*-axis of this plot shows the logarithm of the partition coefficient for the transfer of the Asn+Gln residues from the interior to the exterior of various classes of proteins (as assessed from probe accessibility calculations). The three points represent water-soluble proteins (*22*), the interfacial region of membrane proteins (this work), and the regions of membrane proteins exposed to the apolar regions of bilayers (this work). The linear relationship between the logarithm of the partition coefficients and $\Delta\Delta G$ (for Asn-to-Val) indicates that there is quantitative as well as qualitative agreement between these mutational results and the statistical studies. The free energy associated with an Asn-to-Val mutation changes from favorable to unfavorable as the environment in which the protein is embedded becomes

increasingly nonpolar. In concert, the partition coefficients change in the expected direction. Thus, the experimental results in micelles should provide a realistic and accurate prediction of the energetics in bilayers as well as biological membranes.

MATERIAL AND METHODS

Peptide Synthesis and Purification. The peptide sequences are shown in Figure 1. Peptides were synthesized on an Applied Biosystems model 433A peptide synthesizer. Fmoc (*N*-9-fluorenylmethyloxycaronyl) 2,4-dimethoxybenzyhydrylamine resin (Applied Biosystems) with a substitution level of 0.67 mmol/g was used on a 0.25 mmol scale. To maintain solvation of the peptide on resin, NMP (*N*-methylpyrrolidinone) with 25% DMSO (dimethyl sulfoxide) was chosen as the solvent. Standard couplings were performed as described in Choma et al. (*10*). Some of the difficult amino acids were coupled twice. After each coupling cycle, the peptide was "capped" by reaction with a 10-fold excess of acetic anhydride and diisopropylethylamine for 2 min to avoid the formation of deletion sequences. The resin was dried under reduced pressure, and the peptide was cleaved from the resin (50 mg/mL) with trifluoroacetic acid (TFA) using 5% water (v/v) and 1% triisopropylsilane (v/v) as scavengers for 2 h at room temperature. N-Terminal fluorescent labels were coupled as described in ref 10. After filtering the mixture to remove the resin, TFA was blown off under a gentle nitrogen stream. The peptide was precipitated with equal amounts of cold ether and hexane. After washing several times, organic solvents were removed under reduced pressure.

The cleavage products were solubilized in 50% trifluoroethanol and 50% HPLC buffer A (99.9% water, 0.1% TFA). The solution was sonicated in order to reach complete solubilization. Peptides were purified by reverse-phase HPLC on a C4 preparative column (Vydac) using a linear gradient at 10 mL min^{-1} of buffer A and buffer B (60% 2-propanol, 30% acetonitrile, 10% water, and 0.1% TFA). The peptide molecular weights were confirmed by MALDI-TOF mass spectrometry (PerSeptive Biosystems), and purity was assessed by analytical HPLC (Hewlett-Packard) on a C4 column with a linear gradient using buffer A and buffer B.

SDS−PAGE. A 12% acrylamide Bis-Tris gel in the MES buffer system (Invitrogen, San Diego, CA) was used. Peptides were solubilized in a 4% LDS sample buffer (Invitrogen, San Diego). The gel was run at room temperature at a constant 200 V. NBD labeled peptide bands (except for unlabeled N14) were stained with Colloidal Blue stain (Invitrogen). Apparent molecular weights of pep1 and N14 peptides were determined from a calibration curve obtained using membrane peptides and proteins (*10*).

Analytical Ultracentrifugation. For better absorbance sensitivity, each of the peptides was modified with NBD (7-nitrobenz-2-oxa-1, 3-diazole) at the N-terminus. Peptide concentrations were obtained by measuring absorbance spectra (NBD: $\epsilon_{458} = 21\,000\ M^{-1}$) of each peptide in ethanol solutions. Samples were prepared by mixing peptide/ethanol stock solutions with the appropriate amount of C14-betaine (*N*-tetradecyl-*N*,*N*-dimethyl-3-ammonio-1-propanesulfonate) detergent stock also in ethanol. The C14-betaine and peptide concentrations were varied to study samples at different peptide/detergent ratios. The samples were dried down under reduced pressure, and buffer was added with the appropriate amount of D_2O added to match the solvent's density to the detergent's density. The buffer was 100 mM Na phosphate pH 7.0 with 13% D_2O. The samples were centrifuged in a Beckman XL-I analytical ultracentrifuge at 40 000, 45 000, and 48 000 rpm or 40, 50, and 60 KRPM depending on the cell centerpiece. After the samples were determined to have reached equilibrium, data obtained from UV measurements were analyzed by curve fitting as described in Gratkowski et al. (*9*).

Data for N7, N14, N7-N21, N7-N14, and A7-N14 were initially analyzed using a monomer−trimer equilibrium (with dissociation K_3), providing an estimate of the free energy of association if it is assumed that the system is fully cooperative (no significant formation of intermediate dimeric species). We examined the sensitivity of the global fit to pK_3 by holding pK_3 to a series of defined values. The quality of the resulting fit was assessed using several different metrics including the sum of the squares of the weighted residuals, the randomness of the residuals (quantified using an autocorrelation function), and the summed squared fractional differences between the input and computed integrated concentrations (a check of the fitted baseline). As shown previously (*19*), although N14 is predominantly trimeric, the data can also be fit acceptably well with 10 to 20% dimer at the midpoint of the transition, introducing a corresponding uncertainty in the value of pK_3. Therefore, a final fit was done using freely varying pK_2 and pK_3 to provide an estimate of the uncertainty in pK_3. Pep1 and N21 associated so weakly that although both monomer−dimer and monomer−trimer models gave reasonable data fits, pK's were very imprecise and therefore provide only upper bounds. On the other hand, N7-N14 associated so strongly that not only could no dimerization be accommodated by the fit, but only a lower bound on pK_3 could be obtained.

Lipid Surface Accessibility of Residues in the TM Region. Ten membrane protein structures were selected from the protein data bank for analysis (PDB id: 1c3w, 1e12, lehk, 1eul, 1f88, 1fx8, 1fum, 1jgj, 1ocr, and 1qla). These structures were chosen for clarity in determining residues that interact with lipids. Water molecules, ligands, and detergent molecules were removed prior to calculations. The accessible surface area of the outer boundary surface was calculated using an envelope that wraps around the whole protein by the "envelope" option of the VOLVL package (*17*, *18*). Two probes sizes, 2.5 and 5.0 Å, were used to model lipid molecules. This calculation ignores all interior surfaces involved in ligand and prosthetic group binding. The accessible area for each atom was then recorded, but only a subset of them were used for calculating the lipid propensity, namely, those atoms in predefined TM regions. The TM region for each protein structure was determined manually. To reduce uncertainty due to the definition of TM boundaries, we discarded the first and last two residues in each TM helix.

SUPPORTING INFORMATION AVAILABLE

Additional experimental data. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Popot, J. L., and Engelman, D. M. (2000) Helical membrane protein folding, stability, and evolution. *Annu. Rev. Biochem. 69*, 881−922.
2. Shai, Y. (1995) Molecular recognition between membrane-spanning polypeptides. *Trends Biochem. Sci. 20*, 460−464.
3. White, S. H., and Wimley, W. C. (1999) Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct. 28*, 319−65.
4. Deber, C. M., Liu, L. P., and Wang, C. (1999) Perspective: peptides as mimics of transmembrane segments in proteins. *J. Pept. Res. 54*, 200−5.
5. Lear, J. D., Gratkowski, H., and DeGrado, W. F. (2001) De novo design, synthesis and characterization of membrane-active peptides. *Biochem. Soc. Trans. 29*, 559−64.
6. Adamian, L., and Liang, J. (2001) Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. *J. Mol. Biol. 311*, 891−907.
7. Adamian, L., and Liang, J. (2002) Interhelical hydrogen bonds and spatial motifs in membrane proteins: polar clamps and serine zippers. *Proteins 47*, 209−18.
8. Eilers, M., Shekar, S. C., Shieh, T., Smith, S. O., and Fleming, P. J. (2000) Internal packing of helical membrane proteins. *Proc. Natl. Acad. Sci. U.S.A. 97*, 5796−801.
9. Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2001) Polar side chains drive the association of model transmembrane peptides. *Proc. Natl. Acad. Sci. U.S.A. 98*, 880−5.
10. Choma, C., Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2000) Asparagine-mediated self-association of a model transmembrane helix. *Nature Struct. Biol. 7*, 161−166.
11. Zhou, F. X., Merianos, H. J., Brunger, A. T., and Engelman, D. M. (2001) Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl. Acad. Sci. U.S.A. 98*, 2250−5.
12. Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T., and Engelman, D. M. (2000) Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nature Struct. Biol. 7*, 154−60.
13. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. A. (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded coiled coil. *Science 254*, 539−544.
14. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) A switch between two-, three-, and four-stranded coiled coils. *Science 262*, 1401−1407.
15. Javadpour, M. M., Eilers, M., Groesbeek, M., and Smith, S. O. (1999) Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association. *Biophys. J. 77*, 1609−18.
16. Eilers, M., Patel, A. B., Liu, W. and Smith, S. O. (2002) Comparison of Helix Interactions in Membrane and Soluble alpha-Bundle Proteins. *Biophys. J. 82*, 2720−36.
17. Edelsbrunner, H., Facello, M., and Liang, J. (1996) On the definition and the construction of pockets in macromolecules. *Pac. Symp. Biocomput.* 272−87.
18. Liang, J., Edelsbrunner, H., Fu, P., Sudhakar, P. V., and Subramaniam, S. (1998) Analytical shape computation of macromolecules: II. Inaccessible cavities in proteins. *Proteins 33*, 18−29.
19. Gratkowski, H., Dai, Q.-H., Wand, A. J., DeGrado, W. F., Lear, J. D. (2002) Cooperativity and specificity of association of a designed transmembrane peptide. *Biophys J. 83*, 1613−9.
20. Acharia, A., Ruvinov, S. B., Gai, J., Moll, J. R., Vinson, C. (2002) A heterodimerizing leucine zipper coiled coil system for examining the specificity of **a** position interactions: amino acids I, V, L, N, A, and K. *Biochemistry 41*, 14122−14131.
21. Wagschal, K., Tripet, B., Lavigne, P., Mant, C., and Hodges, R. S. (1999) *Protein Sci. 8*, 2312−29.
22. Miller, S., Janin, J., Lesk, A. M., Chothia, C. (1987) Interior and Surface of Monomeric Proteins *J. Mol. Biol. 196*, 641−656.

BI020573J